ABC SMC for dynamical systems

Tina Toni, Michael Stumpf

Theoretical Systems Biology Group, Imperial College London, UK

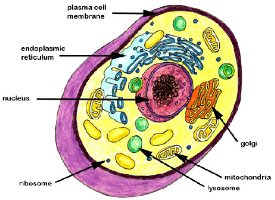
ABC in Paris, 26/06/2009

Outline

Motivation

Complex biological systems

- Models often ODE or stochastic master equations
- High dimensional parameter space
- Time course, non-equidistant, missing data
- Complex likelihood surface



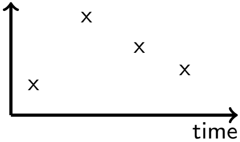
Interested in

- Characterization of distributions over parameters rather than point estimates
- Parameter sensitivity
- Hypothesis testing: nested / non nested models
- Computational efficiency

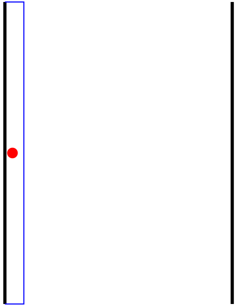
ABC framework for dynamical systems

Prior
$P(\theta)$

Posterior
$P(\theta|D)$

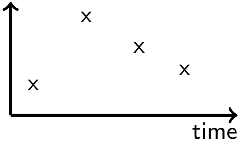


ABC framework for dynamical systems



Prior
$P(\theta)$

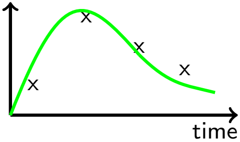
Posterior
$P(\theta|D)$



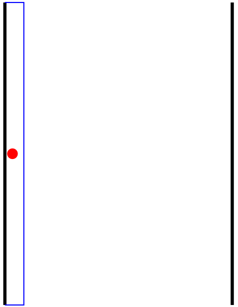
ABC framework for dynamical systems

Prior
$P(\theta)$

Posterior
$P(\theta|D)$

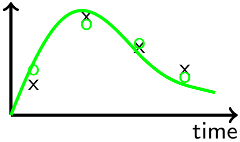


ABC framework for dynamical systems



Prior
$P(\theta)$

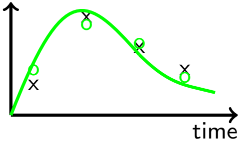
Posterior
$P(\theta|D)$



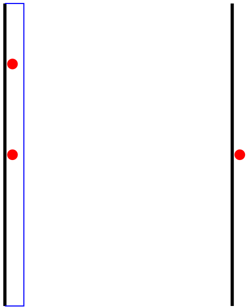
ABC framework for dynamical systems

Prior
$P(\theta)$

Posterior
$P(\theta|D)$

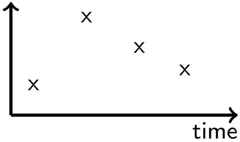


ABC framework for dynamical systems

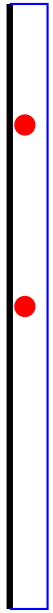


Prior
$P(\theta)$

Posterior
$P(\theta|D)$

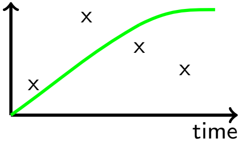


ABC framework for dynamical systems

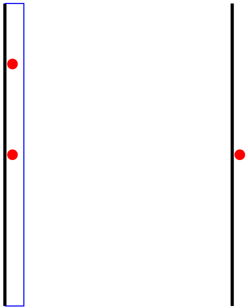


Prior
$P(\theta)$

Posterior
$P(\theta|D)$

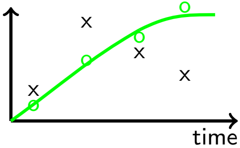


ABC framework for dynamical systems

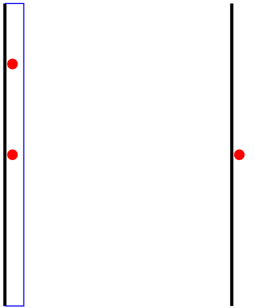


Prior
$P(\theta)$

Posterior
$P(\theta|D)$

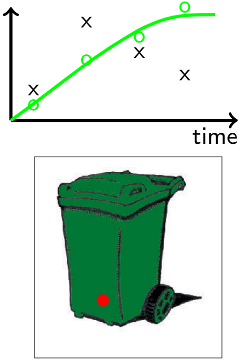


ABC framework for dynamical systems

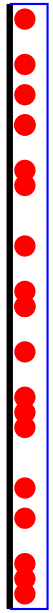


Prior
$P(\theta)$

Posterior
$P(\theta|D)$

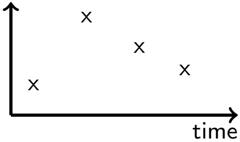


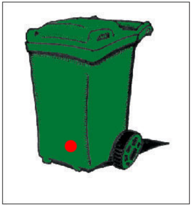
ABC framework for dynamical systems



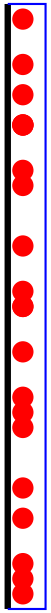
Prior
$P(\theta)$

Posterior
$P(\theta|D)$



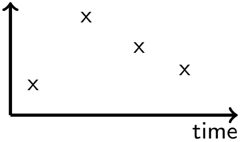


ABC framework for dynamical systems

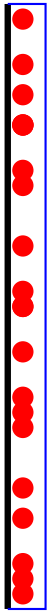


Prior
$P(\theta)$

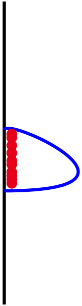
Posterior
$P(\theta|D)$



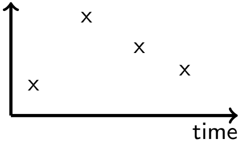
ABC framework for dynamical systems



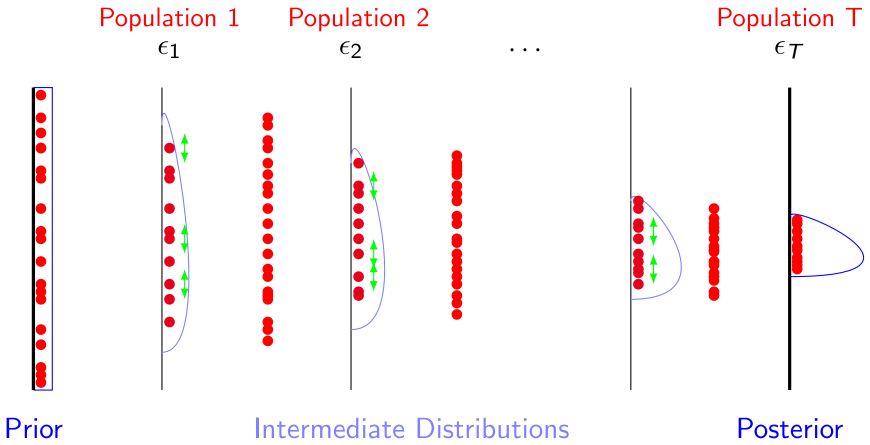
Prior
$P(\theta)$



Posterior
$P(\theta|D)$



ABC SMC for estimation of $P(\theta|D)$



(Sisson et al., 2007)

Weights

Sisson *et al.* (2007)

$$w_t^{(i)} = \frac{\pi(\theta_t^{(i)}) L_{t-1}(\theta^* | \theta_t^{(i)})}{\pi(\theta^*) K_t(\theta_t^{(i)} | \theta^*)} \propto 1 \quad \text{when } L_{t-1} = K_t$$

Instead:

$$\begin{aligned} w_t(\theta_t) &= \frac{\pi_t(\theta_t)}{\eta_t(\theta_t)} \\ \eta_t(\theta_t) &= 1\left(\pi(\theta_t) > 0\right) 1\left(dist > \epsilon_t\right) \int \pi_{t-1}(\theta_{t-1}) K_t(\theta_t | \theta_{t-1}) d\theta_{t-1} \\ w_t^{(i)} &= \frac{\pi(\theta_t^{(i)})}{\sum_{j=1}^{N} w_{t-1}^{(j)} K_t(\theta_t^{(i)} | \theta_{t-1}^{(j)})} \end{aligned}$$

JOURNAL OF THE ROYAL SOCIETY Interface

J. R. Soc. Interface (2009) **6**, 187–202
doi:10.1098/rsif.2008.0172
Published online 9 July 2008

Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems

Tina Toni[1,2,*], David Welch[3,†], Natalja Strelkowa[4], Andreas Ipsen[5] and Michael P. H. Stumpf[1,2,*]

Related

- S. A. Sisson, Y. Fan and M. M. Tanaka. A note on backward kernel choice for sequential Monte Carlo without likelihoods. (S. Sisson's webpage)
- Beaumont MA, Cornuet JM, Marin JM and Robert CP. Adaptive approximate Bayesian computation. (arXiv)

Model selection: $P(M|D) = ?$

1. Marginal likelihoods

- For each model separately estimate $P(D|M)$:

$$P(D|M) = \int_\theta P(D|\theta, M)P(\theta|M)d\theta$$

- Then

$$P(M|D) = \frac{P(D|M)P(M)}{\sum_{M'} P(D|M')P(M')}.$$

2. Joint space

- Include model M as an extra parameter: $(M, \theta^{(1)}, \dots, \theta^{(M)})$
- ABC SMC $\rightarrow P(M, \theta^{(1)}, \dots, \theta^{(M)}|D)$
- Marginalize $\rightarrow P(M|D)$

Example: ABC rejection

1. Marginal likelihoods

- $M = 1$, N_1 particles:
  $P(D|M=1) = \frac{\#\text{accepted}_1}{N_1}$
- $M = 2$, N_2 particles:
  $P(D|M=2) = \frac{\#\text{accepted}_2}{N_2}$
- $\ldots$

$P(M = i|D) =$

$$\frac{\frac{\#\text{accepted}_i}{N_i} P(M=i)}{\sum_j \frac{\#\text{accepted}_j}{N_j} P(M=j)}$$

(Wilkinson, PhD Thesis)

2. Joint space

- Propose M.
- Propose $\theta^{(M)}$.
- Accept/reject $(M, \theta^{(M)})$.

$P(M = i|D) =$

$$\frac{\#\text{accepted particles with } M = i}{\#\text{all accepted particles}}$$

(Grelaud et al., ArXiv)

How can we use ideas from SMC in these approaches?

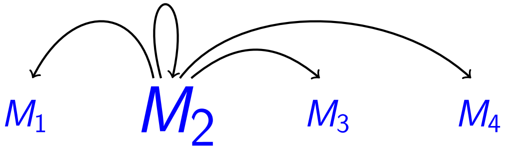
Model selection on a joint space

M_1 M_2 M_3 M_4

Model selection on a joint space

M_1 M_2 M_3 M_4 | M^*

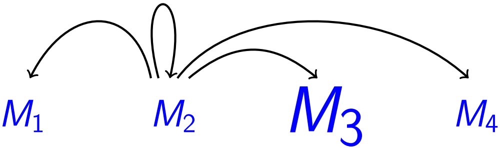
Model selection on a joint space



$$M^*$$

$$M^{**} \sim KM(M|M^*)$$

Model selection on a joint space



$$M^*$$

$$M^{**} \sim KM(M|M^*)$$

Model selection on a joint space

(M_3, θ_3)
(M_3, θ_7)
(M_3, θ_6) (M_3, θ_2)
(M_3, θ_8) (M_3, θ_5) (M_3, θ_1)
(M_3, θ_4)
(M_3, θ_9)

$$M^*$$

$$M^{**} \sim KM(M|M^*)$$

$$\theta^*$$

Model selection on a joint space

(M_3, θ_3)

(M_3, θ_7)

(M_3, θ_6) (M_3, θ_2)

(M_3, θ_8) (M_3, θ_5) (M_3, θ_1)

(M_3, θ_4)

(M_3, θ_9)

M^*

$M^{**} \sim KM(M|M^*)$

θ^*

Model selection on a joint space

(M_3, θ_3)
(M_3, θ_7)
(M_3, θ_6) (M_3, θ_2)
(M_3, θ_8) (M_3, θ_5) (M_3, θ_1)
(M_3, θ_4)
(M_3, θ_9)

$$M^*$$

$$M^{**} \sim KM(M|M^*)$$

$$\theta^*$$

$$\theta^{**} \sim KP(\theta|\theta^*)$$

Model selection on a joint space

$$(M^{**}, \theta^{**})$$

M^*

$M^{**} \sim KM(M|M^*)$

θ^*

$\theta^{**} \sim KP(\theta|\theta^*)$

accept / reject

Model selection on a joint space

$$w(M^{**}, \theta^{**})$$

M^*

$M^{**} \sim KM(M|M^*)$

θ^*

$\theta^{**} \sim KP(\theta|\theta^*)$

accept / reject

calculate w

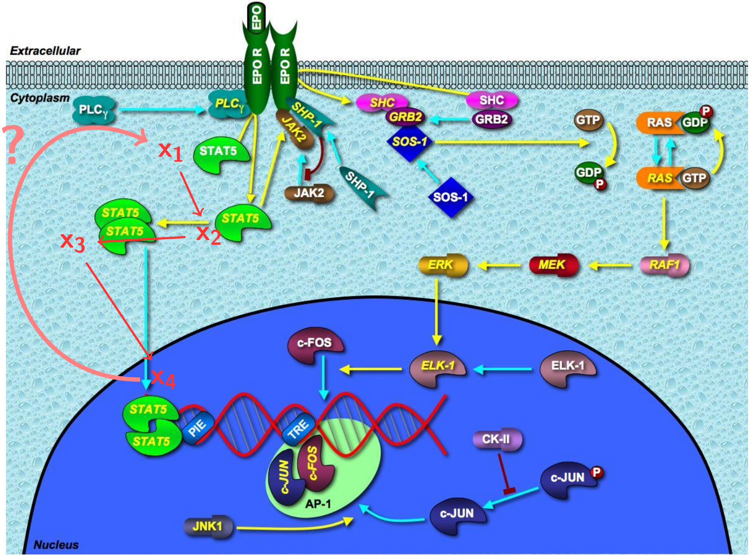
Model selection on joint space: Weight calculation

$$w_t(m^{**}, \theta^{**})$$

$$=$$

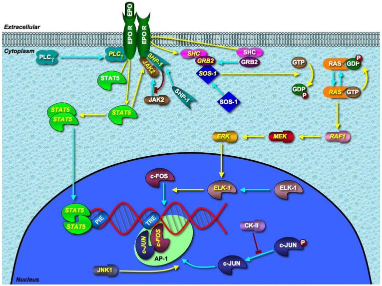
$$\frac{\pi(m^{**}, \theta^{**})}{\underbrace{\sum_{j=1}^{\mathcal{M}} P_{t-1}^{(j)} KM_t(m^{**}|m_{t-1}^{(j)})}_{\text{model perturbation}} \underbrace{\sum_{k;m_{t-1}=m^{**}} \frac{w_{t-1}^{(k)}}{\sum_{l;m_{t-1}=m^{**}} w_{t-1}^{(l)}} KP_{t,m^{**}}(\theta^{**}|\theta_{t-1}^{(k)})}_{\text{parameter perturbation}}}$$

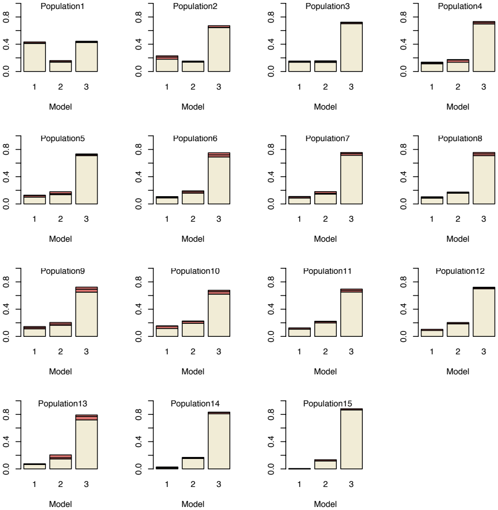
Model selection: JAK-STAT signalling pathway



(adapted from Biocarta)

Model selection: JAK-STAT signalling pathway





Model selection: Phosphorylation dynamics

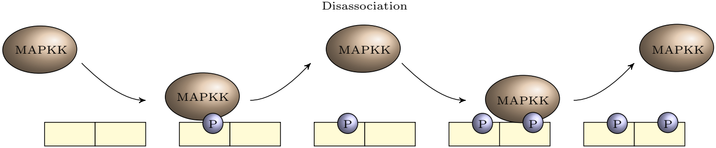


Figure 1: Distributive phosphorylation.

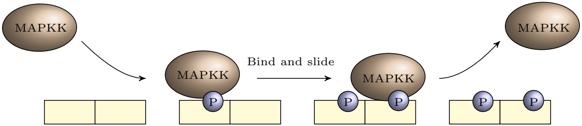
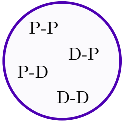
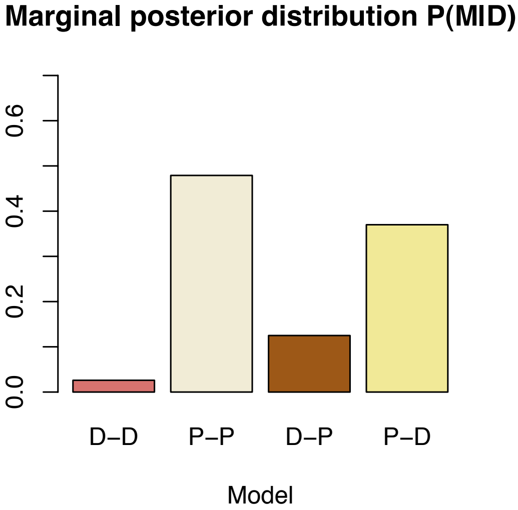


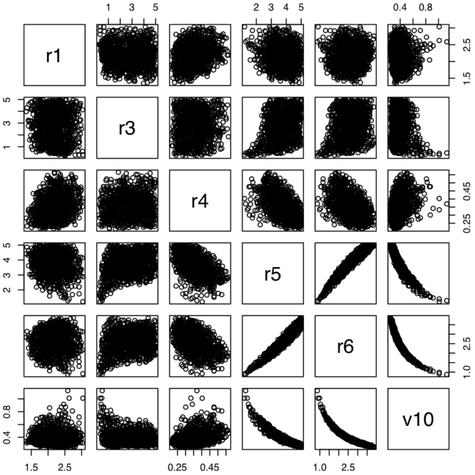
Figure 2: Processive phosphorylation.

Model selection: Phosphorylation dynamics





Parameter estimation: Sensitivity, Identifiability



Pros and Cons

Advantages

- Missing data, non-equidistant time intervals
- Computational improvement over ABC REJ
- Parameter sensitivity for free

Disdvantages

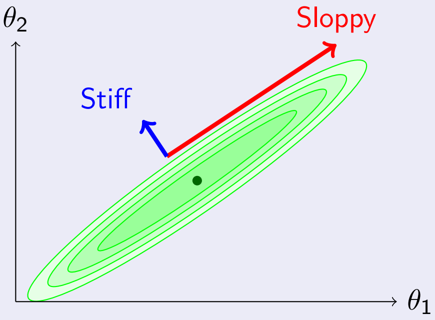
- Computational cost $\longrightarrow$ parallelization
- User input: tolerance schedule $\longrightarrow$ Doucet & Jasra
- Sensitivity to changes in ϵ_T $\longrightarrow$ regression adjustment Leuenberger ($11^{00} - 11^{30}$), Blum ($15^{00} - 15^{30}$), Francois ($17^{00} - 17^{30}$)

Perturbation kernels

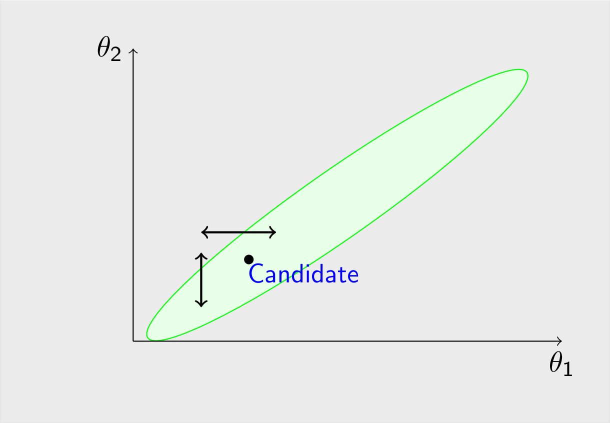
Component-wise

- Gaussian
- Uniform

Parameters in systems biology

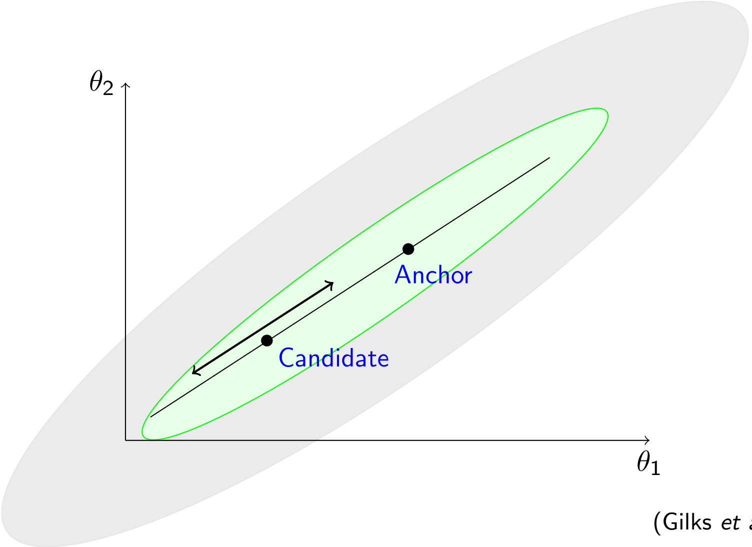


Component-wise perturbation



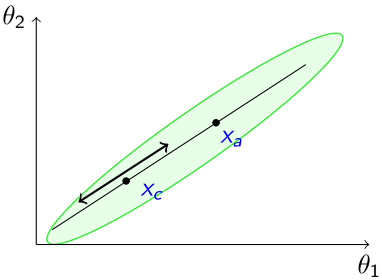
Idea: Perturbation kernels from evolutionary computation and snookering.

Snooker perturbation



(Gilks *et al.*, 1992)

Snooker perturbation



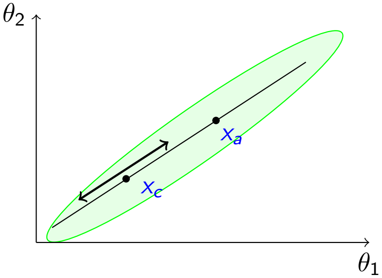
$$x'_c = x_c + r\frac{x_a - x_c}{||x_a - x_c||}$$

$$r \sim N(0, ||x_a - x_c||)$$

$$r \sim U(-1, 1)||x_a - x_c||$$

$$w_t(x'_c) = \frac{\pi(x'_c)}{\sum \frac{w_{x_c} w_{x_a}}{1 - w_{x_c}} K_t(x'_c | x_a, x_c)}$$

Snooker perturbation



$$x'_c = x_c + r\frac{x_a - x_c}{||x_a - x_c||}$$

$$r \sim N(0, ||x_a - x_c||)$$

$$r \sim U(-1,1)||x_a - x_c||$$

$$w_t(x'_c) = \frac{\pi(x'_c)}{\sum \frac{w_{x_c} w_{x_a}}{1-w_{x_c}} K_t(x'_c|x_a, x_c)}$$

Challenge: How to avoid reduced particle diversity?

Crossover

$$\begin{array}{ccc|cc} (a_1 & a_2 & a_3 & a_4 & a_5) \\ (b_1 & b_2 & b_3 & b_4 & b_5) \end{array} \longleftarrow \text{Parent chromosomes}$$

$$\overline{}$$

$$(a_1 \ a_2 \ a_3 \ b_4 \ b_5) \qquad \text{Candidate particle}$$

Crossover

$$\begin{array}{l} (a_1\ a_2\ a_3 \,|\, a_4\ a_5) = x_{t-1}^{(i)} \\ (b_1\ b_2\ b_3 \,|\, b_4\ b_5) = x_{t-1}^{(j)} \end{array} \longleftarrow \text{Parent chromosomes}$$

$$\overline{}$$

$$(a_1\ a_2\ a_3\ b_4\ b_5) = x_t \qquad \text{Candidate particle}$$

$$w(x_t) = \pi(x_t)/\Big(\sum_i \big[\overbrace{(1-\rho)w_{t-1}^{(i)}K_t(x_t|x_{t-1}^{(i)})}^{\text{perturbation}} + \underbrace{\rho\ \frac{1}{N(p-1)}\sum_j w_{t-1}^{(i)}w_{t-1}^{(j)}\ 1(x_t \text{ is recombination of } x_{t-1}^{(i)}, x_{t-1}^{(j)})}_{\text{recombination}}\big]\Big)$$

ρ : probability of recombination

p : particle length

Summary

- ABC SMC for parameter estimation and model selection of dynamical systems.
- Challenges connected to posterior parameter distributions of biological systems:
  - Bayesian framework suited to study parameter sensitivity
  - Sloppiness
  - Perturbation kernels should exploit our knowledge about posterior distributions

Acknowledgements

- David Welch (Penn State University)
- Mark Beaumont (University of Reading)
- Yu-Ichi Ozaki, Shinya Kuroda (University of Tokyo)
- David Balding, Ajay Jasra (Imperial College London)
- Paul Kirk